

Elements of Nonlinear Quantum Mechanics (II): Triple bracket generalization of quantum mechanics

Marek Czachor*

*Pracownia Dielektryków i Półprzewodników Organicznych
Wydział Fizyki Technicznej i Matematyki Stosowanej
Politechnika Gdańska
ul. Narutowicza 11/12, 08-952 Gdańsk, Poland*

An extension of quantum mechanics to a generalized Nambu dynamics leads to a new version of nonlinear quantum mechanics. The time evolution of states is given in here by a triple bracket generalization of the Liouville-von Neumann equation, where one of the generators is an average energy, and the other is a measure of entropy. A nonlinear evolution can occur only for mixed states, and for systems that are described by Rényi α -entropies with $\alpha \neq 2$. The case $\alpha = 2$ corresponds to ordinary, linear quantum mechanics. Since $\alpha = 2$ entropy is the only entropy characterizing systems which cannot gain information, the nonlinear dynamics corresponds to “observers”, that is, systems that *can* gain information. The new formulation of nonlinear quantum mechanics is free from difficulties found in earlier attempts. The connection of linearity with possibilities of gaining information is in a striking agreement with the ideas of Wigner formulated in his paradox of a friend.

I. INTRODUCTION

In the first part of this article [1] I have described the fundamental theoretical difficulties of nonlinear quantum mechanics (NLQM) based on a nonlinear Schrödinger equation. In the present paper I will present a different generalization of quantum mechanics (QM) where the nonlinear evolution never occurs for *pure* states (the Schrödinger equation is hence always linear in this framework) but, instead, may appear, under some circumstances, for *mixed* states. We will see that the new approach will be free from the difficulties discussed in Ref. [1].

The proposed generalization is based on the idea of rewriting the Liouville-von Neumann equation in a triple bracket form, introduced by Białynicki-Birula and Morrison [2]. The triple bracket is an infinite dimensional analog of the Nambu bracket [3] where, as opposed to the structure constants ϵ_{klm} of the rotation algebra appearing in the original Nambu bracket, the structure constants correspond to some infinite-dimensional Lie algebra. In the original Nambu paper an evolution of a physical system (a rigid rotator) is generated by two “Hamiltonian functions”, the energy H and J , where the latter is the Casimir of $so(3)$ (squared angular momentum). The metric tensor used for constructing the Casimir is, as usual, the one related to the Killing form [4]. In the triple bracket formulation of QM the analog of J is the Casimir $S = 1/2\text{Tr}(\rho^2)$ which also can be written as $g^{ab}\rho_a\rho_b$ although, as we shall see later, the metric g^{ab} is no longer given by the Killing-Cartan tensor (which does not exist in this case). The Casimir S was termed in Ref. [2] the *entropy*. It will be argued below that the assignment of the name “entropy” to S should not be regarded as accidental, but as a reflection of a deeper principle relating dynamics with information.

The fact that some sort of such a relationship should be present in QM follows already from the Copenhagen interpretation of a measurement (reduction of a state vector), but my approach will be essentially different and closer in spirit to Wigner’s paradox of a friend [5]. Let me recall that Wigner, in order to solve the paradox, concluded that a conscious observation must be accompanied by a *nonlinear* evolution in the space of the observer’s states. Even though the argumentation of Wigner looks convincing, it seems that standard quantum theories do not leave room for a physical principle of that kind. It is surprising that the triple bracket formalism *does* lead quite naturally to this phenomenon, if we seriously treat the intuitions of Białynicki-Birula and Morrison that S is a measure of quantum entropy.

Putting things more modestly, one can say that the results of this paper, even if their interpretation will turn out inadequate, show that the structure of quantum dynamics may be a part of a more general, nonlinear framework.

The structure of the paper is the following. In Sec. II I discuss various measures of information (entropies) and their possible relationship with quantum mechanics. With this background we will be able to understand why $\text{Tr}(\rho^2)$ is a natural measure of information characterizing systems that *cannot gain information* (and, accordingly, for which the paradox of a friend cannot be formulated). In Sec. III, I present a triple bracket formulation of QM and introduce the composite index form of the equation of motion. In Sec. IV I will discuss an extension of the formalism to nonlinear theories in which the second “Hamiltonian function”, here interpreted as a measure of entropy, is represented by

Casimirs of order higher than 2. Such Casimirs correspond quite naturally to higher order α -entropies discussed in Sec. II. The main technical results of this section are theorems on non-existence of “faster-than-light telegraphs” and on conservation of positivity of ρ_t by the nonlinear evolution equation resulting from the triple bracket formalism. It is also shown that homogeneity preserving generalizations of α -entropies lead to *linear* evolution of *pure* states, and that the $\alpha \rightarrow 1$ limit of α -entropies (the Shannon limit) can be regarded as a kind of classical limit for the generalized QM. Finally, I discuss various possibilities of describing composite systems that consist of subsystems described by different entropies. The paper is concluded with a remark on the complementarity principle in nonlinear QM and an explicitly relativistic formulation is given in the Appendix.

II. MEASURES OF INFORMATION AND QUANTUM MECHANICS

A *logarithmic measure* of information was introduced by R. V. Hartley in 1928 [6]. According to him, to characterize an element of a set of size N we need $\log_2 N$ units of information. It follows that a unit of information (1 bit) is the amount of information necessary for a characterization of a pair. Of course, one can choose also other units such that the unit is the amount of information necessary for a characterization of a set with $0 < k \in \mathbf{N}$ elements, or even with $0 < r \in \mathbf{R}$ elements in average. The respective measures of information in arbitrary units a are $\log_a N$. The most important feature of the logarithmic information measure is its additivity: If a set E is a disjoint union of M N -tuples E_1, \dots, E_M , then we can specify an element of this MN -element set E in two steps: First we need $\log_a M$ units of information to describe which E_k of the sets E_1, \dots, E_M contains the element, then we need $\log_a N$ further units to tell which element of this E_k is the considered one. The information necessary for a characterization of an element of E is the sum of the partial informations: $\log_a MN = \log_a M + \log_a N$. Next step in the development of the measures of information was done independently by C. E. Shannon [7] and N. Wiener [8] in 1948 who derived a formula analogous to Boltzman’s entropy. Their formula has the following heuristic motivation. Let E be the disjoint union of the sets E_1, \dots, E_n having N_1, \dots, N_n elements respectively ($\sum_{k=1}^n N_k = N$). Let us suppose that we are interested only in knowing the subset E_k . (This is typical for classical statistical problems in physics: Statistical quantities depend on classes of microscopic conditions and not on single microscopic properties.) The information characterizing an element of E consists of two parts: The first specifies the subset E_k containing this particular element and the second locates it within E_k . The amount of the second piece of information is, by Hartley formula, $\log_a N_k$ thus depends on the index k . On the other hand, to specify an element of E we need $\log_a N$ units of information. The amount necessary for the specification of the set E_k is therefore

$$I_k = \log_a N - \log_a N_k = \log_a \frac{N}{N_k} = \log_a \frac{1}{p_k}. \quad (1)$$

It follows that the amount of information *received by learning* that a single event of probability p took place equals

$$I(p) = \log_a \frac{1}{p}. \quad (2)$$

In statistical situations measured quantities correspond to averages of random variables. Therefore the average information is

$$I = \sum_k p_k \log_a \frac{1}{p_k}. \quad (3)$$

This is the Shannon’s formula and I is called the *entropy* of the probability distribution $\{p_1, \dots, p_n\}$. If all the probabilities are equal $1/N$ then the Shannon’s formula is equal to the Hartley’s one. The mean we have applied is the so-called linear mean. Rényi observed that there exist information theoretic problems where the measures of information are those obtained by more general ways of averaging — the Kolmogorov–Nagumo function approach [9]. Let φ be a monotonic function on real numbers. The Kolmogorov–Nagumo average information can be defined by means of φ as

$$I = \varphi^{-1} \left(\sum_k p_k \varphi \left(\log_a \frac{1}{p_k} \right) \right). \quad (4)$$

If the generalized information measure is to satisfy the postulate of additivity, φ must be a linear or exponential function. The linear function corresponds to Shannon’s information. The exponential functions provide a large class

of new measures of information. Consider a function $\varphi(x) = a^{(1-\alpha)x}$. We can always choose the units of information in such a way that

$$I = \varphi^{-1} \left(\sum_k p_k \varphi \left(\log_a \frac{1}{p_k} \right) \right) = \frac{1}{1-\alpha} \log_a \left(\sum_k p_k^\alpha \right) = \log_a \left(\left(\sum_k p_k^\alpha \right)^{1/(1-\alpha)} \right). \quad (5)$$

For $p_k = 1/N$ we obtain again the Hartley formula. Formula (5) describes Rényi's α -entropy which, from now on, will be denoted $I_\alpha(\mathcal{P})$, where \mathcal{P} denotes the probability distribution. We see that the essential part of the definition is played by

$$I_\alpha^*(\mathcal{P}) = a^{I_\alpha(\mathcal{P})} = \left(\sum_k p_k^\alpha \right)^{1/(1-\alpha)} \quad (6)$$

which is independent of the choice of the unit a . To distinguish between α -entropy and $I_\alpha^*(\mathcal{P})$ we shall call the latter α^* -entropy (* will remind us that this quantity is multiplicative in opposition to the additivity of $I_\alpha(\mathcal{P})$). (The observation that what is in fact informationally fundamental in $I_\alpha(\mathcal{P})$ is $I_\alpha^*(\mathcal{P})$ is strengthened by Daróczy's definition of *entropy of order α* [10] defined as

$$(2^{1-\alpha} - 1)^{-1} \left(\sum_k p_k^\alpha - 1 \right). \quad (7)$$

This expression possesses many ordinary properties of the entropy and in the limit $\alpha \rightarrow 1$ becomes, the so-called Shannon's information function.)

The limit $\alpha \rightarrow 1$ is interesting also for α -entropies. It can be shown that $I_1 = \lim_{\alpha \rightarrow 1} I_\alpha$ equals Shannon's entropy.

$I_\alpha(\mathcal{P})$ is a monotonic, decreasing function of α . For negative α $I_\alpha(\mathcal{P})$ tends to infinity if one of p_k tends to zero. This property excludes $\alpha < 0$ because adding a new event of probability 0 to a probability distribution, what does not change the probability distribution, turns $I_\alpha(\mathcal{P})$ into infinity.

A fundamental notion in information theory is the *gain of information*. Consider an experiment whose results are A_1, \dots, A_n having probabilities $p_k = P(A = A_k)$. We observe an event B related to the experiment and obtain a result $B = B_l$. Now the conditional probabilities are $p_{kl} = P(A = A_k | B = B_l)$. Consider now a system (an "observer") whose information is measured by some α -entropy. How much information about the random variable A has he received by observation of $B = B_l$? The amount of information he would have obtained by observing $A = A_k$ would be equal to

$$\log_a \frac{1}{p_k} \quad (8)$$

if he had not measured B . After having observed $B = B_l$ the amount of information he would have obtained by observing $A = A_k$ would be

$$\log_a \frac{1}{p_{kl}}. \quad (9)$$

It follows that the measurement of $B = B_l$ has given him already

$$\log_a \frac{1}{p_k} - \log_a \frac{1}{p_{kl}} = \log_a \frac{p_k}{p_{kl}} \quad (10)$$

units of information about A . The expression (10) is called the decrease of uncertainty about $A = A_k$ by observing $B = B_l$. We define the gain of information about A , obtained when the probability distribution $\{p_k\}$ is replaced by $\{p_{kl}\}$, by

$$\varphi^{-1} \left(\sum_k p_{kl} \varphi \left(\log_a \frac{p_k}{p_{kl}} \right) \right) = \frac{1}{1-\alpha} \log_a \left(\sum_k \frac{p_k^{2-\alpha}}{p_{kl}^{1-\alpha}} \right). \quad (11)$$

If we define the increase of the uncertainty by minus decrease of uncertainty we can calculate the average "loss of information" defined by

$$\varphi^{-1} \left(\sum_k p_{kl} \varphi \left(\log_a \frac{p_{kl}}{p_k} \right) \right) = \frac{1}{1-\alpha} \log_a \left(\sum_k \frac{p_{kl}^\alpha}{p_k^{\alpha-1}} \right). \quad (12)$$

For Shannon's entropy the gain is minus the loss. For α -entropies the two concepts are inequivalent.

The gain of information defined by (11) for $\alpha > 2$ has the same pathological properties as I_α for $\alpha < 0$ so, it seems, cannot be consistently applied unless we restrict $0 \leq \alpha \leq 2$. This is the reason why Rényi defined the gain of information as minus the loss, although such a definition is less natural. From the viewpoint of our quantum mechanical applications the situation is not so clear, however, and the following argument shows that $\alpha = 2$ is a natural value limiting α -s from above.

When we speak about information, what we have in mind is not the subjective "information" possessed by a particular, animate observer. In reality the information contained in an observation is a quantity independent of the fact whether it does or does not reach the perception of the observer (be it a man, some registering device, a computer, or some other physical system). On the other hand, different kinds of entropies introduced above may be characteristic for different systems. The entropy (information) is objective in the same sense as probability, and in the same sense it is reasonable to expect that there are classical and quantum informations, as there are classical and quantum probabilities.

The procedure leading to the notion of the decrease of uncertainty assumes implicitly that after each measurement of a random variable, here B , one can always proceed further in getting information about A , and that the procedure terminates when we know everything about the state of the system. In classical world this final state of knowledge means no uncertainties. Therefore, classically, if there is some lack of knowledge about a system, then there exists, in principle, a possibility of *gaining information*. Putting it more formally, we can say that an information characterizing a *classical* system should allow for different gains of information in different situations. The quantum mechanical no-hidden-variables postulate means that the probabilistic description of a quantum system does not follow from our lack of knowledge about the system. This suggests that a quantum information, characterizing a quantum system, might be of such a kind that its corresponding gain of information is zero under all circumstances. It is tempting to develop this hypothesis a little and find whether a measure of information possessing this property exists.

The Shannon's information gain is given by

$$-\sum_k p_{kl} \log_a \frac{p_{kl}}{p_k} \quad (13)$$

and vanishes only if A and B are independent. So this case can be excluded because we want the gain of information to be 0 for all probability distributions (this excludes also the von Neumann entropy). For α -entropies we find that the vanishing of (11) implies

$$\sum_k \frac{p_{kl}^{2-\alpha}}{p_k^{1-\alpha}} = \sum_k p_k (p_{kl}^{2-\alpha} p_k^{\alpha-2}) = 1 \quad (14)$$

which can hold for all p_k and p_{kl} if and only if $\alpha = 2$. It follows that the only candidate for the quantum entropy is the Rényi's 2-entropy which reads

$$-\log_a \left(\sum_k p_k^2 \right). \quad (15)$$

Expressing the probabilities by means of a density matrix and choosing the unit of information with $a = e$ we obtain

$$I_2[\rho] = -\ln \text{Tr}(\rho^2). \quad (16)$$

This kind of entropy is sometimes considered as an alternative to von Neumann's entropy [11]. Our reasoning, based on the assumption that an *ordinary quantum* system should not have a possibility of gaining information, selects this entropy in a unique way. It is clear, from the perspective of the Wigner's paradox of a friend, that *observers*, who can gain information, should be described by $\alpha \neq 2$ -entropies.

III. POISSONIAN FORMULATION OF QUANTUM MECHANICS

A departure point for the discussed generalization of linear QM is the observation that quantum theory can be regarded as a particular *classical* infinite dimensional Hamiltonian, Poissonian or Nambu-like theory.

Let \mathcal{H} be a Hilbert space. Consider the Hamilton equations

$$\omega^{AA'}(\alpha, \alpha') \frac{d\psi_A(\alpha)}{d\tau} = \frac{\delta H}{\delta \psi_{A'}^*(\alpha')} \quad (17)$$

and c.c., where the bars denote complex conjugations and the conventions concerning primed and unprimed indices are assumed like in the spinor abstract index calculus [12]. The summation convention is as follows: We sum over repeated Roman indices and integrate over repeated Greek ones. The integration is with respect to some invariant, or quasi-invariant measure on a finite dimensional manifold (mass hyperboloid, spacelike hyperplane in the Minkowski space, etc.). The symbol of the “proper time” derivative describes a differentiation with respect to a suitable foliation of space-time (Minkowskian spacelike, or Galilean $t = \text{const}$ hyperplanes, etc., see Appendix). In Hilbertian formulation of QM the “symplectic form” is given by the delta distribution

$$\omega^{AA'}(\alpha, \alpha') := i\delta^{AA'}\delta(\alpha, \alpha') =: \omega^{AA'}\delta(\alpha, \alpha') \quad (18)$$

where $\delta^{AB'} = \delta_{AB'} = 1$ if $A = B'$ and 0 for $A \neq B'$ in the nonrelativistic QM. For the Dirac equation $\delta^{AB'}$ and $\delta_{AB'}$ can be represented by the Dirac matrix γ_0 , and the Dirac delta function must correspond to the choice of the spacelike hyperplane. In the projective space formulation the symplectic form corresponds to the Fubini-Study metric. The inverse of $\omega^{AA'}(\alpha, \alpha')$ is

$$I_{AA'}(\alpha, \alpha') := -i\delta_{AA'}\delta(\alpha, \alpha') =: I_{AA'}\delta(\alpha, \alpha') \quad (19)$$

where by the inverse we understand that

$$\omega^{AA'}(\alpha, \alpha')I_{BA'}(\beta, \alpha') = \delta_B^A\delta(\alpha, \beta) \quad (20)$$

$$\omega^{AA'}(\alpha, \alpha')I_{AB'}(\alpha, \beta') = \delta_B^A\delta(\alpha', \beta'). \quad (21)$$

Accordingly, the form of the Hamilton equations we shall use is

$$\frac{d\psi_A(\alpha)}{d\tau} = I_{AA'}\frac{\delta H}{\delta\psi_{A'}^*(\alpha)} \quad (22)$$

and c.c. (22) describes a quantum evolution of pure states. All observables of the linear theory depend on $|\psi\rangle$ and $\langle\psi|$ via the density matrix $\rho = |\psi\rangle\langle\psi|$. Let F and G be two such observables, that is $F[\psi, \psi^*] = F[\rho]$ and $G[\psi, \psi^*] = G[\rho]$. The Poisson bracket resulting from the Hamilton equations is

$$\{F, G\} = I_{AA'}\left(\frac{\delta F}{\delta\psi_A(\alpha)}\frac{\delta G}{\delta\psi_{A'}^*(\alpha)} - \frac{\delta G}{\delta\psi_A(\alpha)}\frac{\delta F}{\delta\psi_{A'}^*(\alpha)}\right). \quad (23)$$

Applying the chain rule to the components of the pure state density matrix

$$\rho_{AA'}(\alpha, \alpha') = \psi_A(\alpha)\psi_{A'}^*(\alpha') \quad (24)$$

we find that

$$\{F, G\} = I_{AA'}\left(\frac{\delta F}{\delta\rho_{AB'}(\alpha, \beta')}\rho_{CB'}(\gamma, \beta')\frac{\delta G}{\delta\rho_{CA'}(\gamma, \alpha)} - (F \leftrightarrow G)\right). \quad (25)$$

So long as the density matrix in (25) is given by (24) the bracket is equivalent to the Poisson bracket (23). Jordan, in a context of the Weinberg’s theory [13] and for a finite dimensional Hilbert space, investigated properties of the bracket (25) with ρ being an arbitrary density matrix. For reasons that will be explained below I will term such a general bracket the Białynicki-Birula–Morrison–Jordan (BBMJ) bracket.

We will now show that (25), for a general ρ , can be written in a form of a generalized Nambu bracket. Let ρ be arbitrary. The BBJM bracket can be rewritten as

$$\{F, G\} = \rho_{AA'}(\alpha, \alpha')\Omega^{AA'}_{BB'CC'}(\alpha, \alpha', \beta, \beta', \gamma, \gamma')\frac{\delta F}{\delta\rho_{BB'}(\beta, \beta')}\frac{\delta G}{\delta\rho_{CC'}(\gamma, \gamma')} \quad (26)$$

with

$$\begin{aligned} \Omega^{AA'}_{BB'CC'}(\alpha, \alpha', \beta, \beta', \gamma, \gamma') &= \delta_C^A\delta_{B'}^{A'}I_{BC'}\delta(\alpha, \gamma)\delta(\alpha', \beta')\delta(\beta, \gamma') \\ &\quad - \delta_B^A\delta_{C'}^{A'}I_{CB'}\delta(\alpha, \beta)\delta(\alpha', \gamma')\delta(\gamma, \beta') \\ &= \Omega^a_{bc} \end{aligned} \quad (27)$$

where, in analogy to the spinor calculus, we have clumped together the respective quadruples of indices into composite ones ($a = (A, A', \alpha, \alpha')$, etc.).

The “structure kernels” Ω_{bc}^a satisfy conditions characteristic for Lie-algebraic structure constants:

$$\Omega_{cb}^a = -\Omega_{bc}^a \quad (28)$$

and

$$\Omega_{bc}^a \Omega_{de}^c + \Omega_{ec}^a \Omega_{bd}^c + \Omega_{dc}^a \Omega_{eb}^c = 0 \quad (29)$$

These two conditions imply the Jacobi identity. The composite index form of the BBMJ bracket

$$\{F, G\} = \rho_a \Omega_{bc}^a \frac{\delta F}{\delta \rho_b} \frac{\delta G}{\delta \rho_c} \quad (30)$$

shows that it takes the same form as the generalized BBM-Nambu bracket written in terms of the Wigner function for a scalar field [2]. As a matter of fact, the BBMJ bracket is simply a different representation of the BBM bracket. The formula (30) looks much the same as the Poisson bracket related to the Kiryloff form on coadjoint representations of Lie groups [14] (such brackets for general structure constants are called the Lie-Poisson brackets (cf. [15])).

It remains to find out how to formulate the explicit *triple bracket* equivalent to (30).

In order to do this we first have to define a “metric tensor” to lower the upper index in the structure kernels (Białynicki-Birula and Morrison avoided this difficulty because the field they considered had no spinor components). The apparently natural guess (the Killing-Cartan metric)

$$g_{ab} = \Omega_{ad}^c \Omega_{bc}^d \quad (31)$$

is incorrect as (31) involves expressions like $\delta(0)$ which are not distributions in the Schwartz sense.

The correct definitions are

$$g_{ab} = -I_{AB'}(\alpha, \beta') I_{BA'}(\beta, \alpha') \quad (32)$$

$$g^{ab} = -\omega^{AB'}(\alpha, \beta') \omega^{BA'}(\beta, \alpha'). \quad (33)$$

The metric tensor is symmetric

$$g_{ab} = g_{ba} \quad (34)$$

and satisfies the invertibility conditions

$$g^{ab} g_{bc} = g_{cb} g^{ba} = \delta_C^A \delta_{C'}^{A'} \delta(\alpha, \gamma) \delta(\alpha', \gamma') =: \delta_c^a. \quad (35)$$

The metric tensor is a useful tool. Consider for example a ρ -independent $F_b = F_{BB'}(\beta, \beta')$. Then

$$F[\rho] = g^{ab} \rho_a F_b = \rho^{B'}_{A'}(\beta', \alpha') F^{A'}_{B'}(\alpha', \beta') = \text{Tr } \rho \hat{F} \quad (36)$$

and we see that linear observables can be naturally expressed with the help of (33). This example is important also as an illustration of the convention concerning lowering and raising of indices. For notice that

$$\frac{\delta F}{\delta \rho^{B'}_{A'}(\beta', \alpha')} = F^{A'}_{B'}(\alpha', \beta') \quad (37)$$

although the staggering of indices like $F_{B'}^{A'}(\beta', \alpha')$ might seem more natural.

The fully covariant form of the structure kernels is

$$\Omega_{abc} = -I_{AB'}(\alpha, \beta') I_{CA'}(\gamma, \alpha') I_{BC'}(\beta, \gamma') + I_{AC'}(\alpha, \gamma') I_{BA'}(\beta, \alpha') I_{CB'}(\gamma, \beta'). \quad (38)$$

One easily verifies that Ω_{abc} is totally antisymmetric.

Following Białynicki-Birula and Morrison let us introduce the functional

$$S_2 = \frac{1}{2} g^{ab} \rho_a \rho_b = \frac{1}{2} \text{Tr } (\rho^2), \quad (39)$$

which is one half of the inverse of Rényi's 2*-entropy.

The BBMJ bracket is now equal to the following triple bracket

$$\{F, G\} = [F, G, S_2] = \Omega_{abc} \frac{\delta F}{\delta \rho_a} \frac{\delta G}{\delta \rho_b} \frac{\delta S_2}{\delta \rho_c}. \quad (40)$$

The antisymmetry of the triple bracket means that S_2 is the Casimir for the BBMJ bracket Lie algebra of observables. Another Casimir is $\text{Tr } \rho$ because $\{\text{Tr } \rho, F\} = 0$ for *any* differentiable F (hence not only linear). The wave functions have been eliminated from the dynamical equations, but the Hilbert space background is implicitly present in the structure kernels and the metric tensor which are defined in terms of ω and I , and in the very notion of the density matrix which acts in the Hilbert space.

Components of the pure state density matrix satisfy

$$\frac{d}{d\tau} \rho_a = \{\rho_a, H\}. \quad (41)$$

which holds also for general density matrices as can be seen from the familiar, operator version of the Liouville–von Neumann equation. It follows that the density matrices form a Poisson manifold, as opposed to state vectors that form a phase space.

IV. NONLINEAR QUANTUM MECHANICS AS A GENERALIZED NAMBU MECHANICS

The generalizations of quantum mechanics considered by Kibble [16] and Weinberg [17] are based on the Hamiltonian framework. The nonlinear evolution is introduced through an extension of the class of admissible Hamiltonian functions. More generally, all canonical transformations are generated by a larger class of functionals on Hilbert or projective spaces. The functionals are a generalization of averages of observable quantities. This fact leads to the fundamental difficulty in constructing a probability interpretation of such theories: The generalized observables do not form an associative algebra which makes impossible a unique definition of powers of observables, the formal counterpart of higher moments of random variables measured in experiments.

The triple bracket form of the Liouville–von Neumann equation shows that the time evolution in linear QM has, in fact, *two* generators: the average energy (Hamiltonian function) and the Casimir S , which measures Rényi's $\alpha = 2$ entropy (or, even more directly, Daróczy entropy of order 2). It is natural to ask what will be changed in the theory if, instead of generalizing the class of admissible Hamiltonian functions, we shall extend the class of entropies. A physical meaning of such an extension would be the one required by Wigner in his paradox of a friend: We extend quantum mechanics to systems that can gain information. The extended theory has a well defined probability interpretation, because the observables are represented by linear operators, provided the scaling by a constant, $\rho \rightarrow \lambda \rho$, is a symmetry of the dynamics. This imposes on the generalized entropies the 2-homogeneity condition: $S(\lambda \rho) = \lambda^2 S(\rho)$.

Only for $S[\rho] = 1/2 \text{Tr}(\rho^2)$ the linear observables are closed under the action of the bracket $\{\cdot, \cdot\}_S := [\cdot, \cdot, S]$. If we extend the class of acceptable S , we have to accept also a somewhat stronger form of the complementarity principle than in linear QM: Observables are always complementary to their (nonvanishing) time derivatives (see Sec. V). We shall begin the discussion of the generalization with the question whether, for general S , the manifold of states is the Poisson manifold.

A. The Jacobi Identity

Let F, G, H and S be arbitrary twice functionally differentiable functionals. We consider the expression

$$\begin{aligned} J &= \{\{F, G\}_S, H\}_S + \{\{H, F\}_S, G\}_S + \{\{G, H\}_S, F\}_S \\ &= \frac{\delta F}{\delta \rho_d} \frac{\delta G}{\delta \rho_e} \frac{\delta^2 S}{\delta \rho_a \delta \rho_f} \frac{\delta H}{\delta \rho_b} \frac{\delta S}{\delta \rho_c} (\Omega_{def} \Omega_{abc} + \Omega_{bdf} \Omega_{aec} + \Omega_{ebf} \Omega_{adc}) \end{aligned} \quad (42)$$

which holds good for any S . $\frac{\delta^2 S}{\delta \rho_a \delta \rho_f} = g^{af}$ for $S = S_2$ and (42) vanishes in virtue of (29). For more general $S = S(f_2[\rho])$ we find

$$\frac{\delta S}{\delta \rho_c} = 2 \frac{\partial S}{\partial f_2} \rho^c \quad (43)$$

$$\frac{\delta^2 S}{\delta \rho_a \delta \rho_f} = 4 \frac{\partial^2 S}{\partial f_2^2} \rho^a \rho^f + 2 \frac{\partial S}{\partial f_2} g^{af}. \quad (44)$$

Inserting these expressions into (42) we obtain

$$J = 8 \frac{\delta F}{\delta \rho_d} \frac{\delta G}{\delta \rho_e} \frac{\partial^2 S}{\partial f_2^2} \rho^a \rho^f \frac{\delta H}{\delta \rho_b} \frac{\partial S}{\partial f_2} \rho^c (\Omega_{def} \Omega_{abc} + \Omega_{bdf} \Omega_{aec} + \Omega_{ebf} \Omega_{adc}) = 0 \quad (45)$$

since $\Omega_{abc} \rho^a \rho^c = 0$. With this choice of S we obtain the dynamics given by

$$\frac{d}{d\tau} \rho_a = \{\rho_a, H\}_{S_2} C[\rho] \quad (46)$$

where $C[\rho] = 2 \frac{\partial S}{\partial f_2} = C(f_2[\rho])$ is an integral of motion, as we shall see later. The only difference with respect to ordinary QM would be in a ρ -dependent rescaling of time, a phenomenon that, in principle, might influence lifetime characteristics of physical processes.

For more general S the question of the Jacobi identity is open, hence we have to accept the possibility that mixed states in the generalized QM do not form a Poisson manifold. This would not be surprising, since in various versions of generalizations of the Nambu mechanics, the Jacobi identity does not hold.

B. Composite Systems in the New Framework

Let the Hilbert space in question and the density matrix of some composite system be $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and

$$\rho_a = \rho_{AA'}(\alpha, \alpha') = \rho_{A_1 A_2 A'_1 A'_2}(\alpha_1, \alpha_2, \alpha'_1, \alpha'_2). \quad (47)$$

The same doubling of indices concerns

$$I_{AA'}(\alpha, \alpha') = -i \delta_{A_1 A'_1} \delta_{A_2 A'_2} \delta(\alpha_1, \alpha'_1) \delta(\alpha_2, \alpha'_2). \quad (48)$$

Reduced density matrices of the two subsystems are

$$\rho_{A_1 A'_1}^I(\alpha_1, \alpha'_1) = \delta^{A_2 A'_2} \delta(\alpha_2, \alpha'_2) \rho_{A_1 A_2 A'_1 A'_2}(\alpha_1, \alpha_2, \alpha'_1, \alpha'_2) \quad (49)$$

$$\rho_{A_2 A'_2}^{II}(\alpha_2, \alpha'_2) = \delta^{A_1 A'_1} \delta(\alpha_1, \alpha'_1) \rho_{A_1 A_2 A'_1 A'_2}(\alpha_1, \alpha_2, \alpha'_1, \alpha'_2) \quad (50)$$

and satisfy

$$\frac{\delta \rho_{A_1 A'_1}^I(\alpha_1, \alpha'_1)}{\delta \rho_{B_1 B_2 B'_1 B'_2}(\beta_1, \beta_2, \beta'_1, \beta'_2)} = \delta_{A_1}^{B_1} \delta^{B_2 B'_2} \delta_{A'_1}^{B'_1} \delta(\beta_1, \alpha_1) \delta(\beta_2, \beta'_2) \delta(\beta'_1, \alpha'_1) \quad (51)$$

and

$$\frac{\delta \rho_{A_2 A'_2}^{II}(\alpha_2, \alpha'_2)}{\delta \rho_{B_1 B_2 B'_1 B'_2}(\beta_1, \beta_2, \beta'_1, \beta'_2)} = \delta_{A_2}^{B_2} \delta^{B_1 B'_1} \delta_{A'_2}^{B'_2} \delta(\beta_2, \alpha_2) \delta(\beta_1, \beta'_1) \delta(\beta'_2, \alpha'_2). \quad (52)$$

The structure kernels for the composite system are

$$\begin{aligned} \Omega_{abc} &= \Omega_{a_1 a_2 b_1 b_2 c_1 c_2} \\ &= -i \left(\delta_{A_1 B'_1} \delta_{C_1 A'_1} \delta_{B_1 C'_1} \delta_{A_2 B'_2} \delta_{C_2 A'_2} \delta_{B_2 C'_2} \times \right. \\ &\quad \delta(\alpha_1, \beta'_1) \delta(\gamma_1, \alpha'_1) \delta(\beta_1, \gamma'_1) \delta(\alpha_2, \beta'_2) \delta(\gamma_2, \alpha'_2) \delta(\beta_2, \gamma'_2) \\ &\quad - \delta_{A_1 C'_1} \delta_{B_1 A'_1} \delta_{C_1 B'_1} \delta_{A_2 C'_2} \delta_{B_2 A'_2} \delta_{C_2 B'_2} \times \\ &\quad \left. \delta(\alpha_1, \gamma'_1) \delta(\beta_1, \alpha'_1) \delta(\gamma_1, \beta'_1) \delta(\alpha_2, \gamma'_2) \delta(\beta_2, \alpha'_2) \delta(\gamma_2, \beta'_2) \right). \end{aligned} \quad (53)$$

The following two results solve generally the question of faster-than-light telegraphs in both Hamiltonian and triple bracket frameworks.

Lemma 1 *Reduced density matrices of the subsystems satisfy*

$$\Omega_{abc} \frac{\delta \rho_{d_1}^I}{\delta \rho_a} \frac{\delta \rho_{d_2}^{II}}{\delta \rho_b} = 0. \quad (54)$$

Proof: It is sufficient to contract (53) with (51) and (52). \square

Theorem 2 *Let $F = F[\rho^I]$ and $G = G[\rho^{II}]$, that is depend on ρ via (49) and (50), then for any S*

$$\{F, G\}_S = 0. \quad (55)$$

Proof: By virtue of the lemma one has

$$0 = \Omega_{abc} \frac{\delta \rho_{d_1}^I}{\delta \rho_a} \frac{\delta \rho_{d_2}^{II}}{\delta \rho_b} \frac{\delta F}{\delta \rho_{d_1}^I} \frac{\delta G}{\delta \rho_{d_2}^{II}} \frac{\delta S}{\delta \rho_c} = \{F, G\}_S. \quad (56)$$

\square

Notice that we have not assumed anything but differentiability not only about S but also about F and G . So, in particular, for arbitrary (nonlinear) observables and $S = S_2$ we obtain the Polchinski-Jordan result for Weinberg's nonlinear QM.

C. Density Matrix Interpretation of Solutions of the Generalized Evolution Equation

One of the essential questions we have to clarify concerns the density matrix interpretation of the solutions of the generalized Liouville-von Neumann equation

$$\frac{d}{d\tau} \rho_a = [\rho_a, H, S]. \quad (57)$$

There is no general *a priori* guarantee that the generalized dynamics will conserve positivity of ρ . The next theorems will give a partial answer to this problem.

In order to attack the question we have to make the language of the S -brackets more readable. Consider the triple bracket $[F, G, H]$ of arbitrary functionals F , G and H . We find that

$$\begin{aligned} i[F, G, H] = & \frac{\delta F}{\delta \rho_{A'}^{B'}(\beta', \alpha')} \frac{\delta G}{\delta \rho_{B'}^{C'}(\gamma', \beta')} \frac{\delta H}{\delta \rho_{C'}^{A'}(\alpha', \gamma')} \\ & - \frac{\delta F}{\delta \rho_{A'}^{C'}(\gamma', \alpha')} \frac{\delta G}{\delta \rho_{B'}^{A'}(\alpha', \beta')} \frac{\delta H}{\delta \rho_{C'}^{B'}(\beta', \gamma')}. \end{aligned} \quad (58)$$

Applying the notation of (37) (where now the “operator” kernels are in general ρ -dependent) we transform (58) into

$$\begin{aligned} F_{B'}^{A'}(\alpha', \beta') G_{C'}^{B'}(\beta', \gamma') H_{A'}^{C'}(\gamma', \alpha') \\ - F_{C'}^{A'}(\alpha', \gamma') G_{A'}^{B'}(\beta', \alpha') H_{B'}^{C'}(\gamma', \beta') = \text{Tr}([\hat{F}, \hat{G}]\hat{H}). \end{aligned} \quad (59)$$

In the last line we have introduced an abbreviated convention based on the assignment to any functional F of an operator

$$\hat{F} = \frac{\delta F}{\delta \rho} \quad (60)$$

which is defined by the kernel form used in (59). For example

$$\rho = \frac{\delta S_2}{\delta \rho}, \quad (61)$$

and

$$\frac{\delta \text{Tr}(\rho^n)}{\delta \rho} = n\rho^{n-1}, \quad (62)$$

the latter being the shortened form of

$$\begin{aligned} \frac{\delta \text{Tr}(\rho^n)}{\delta \rho_{AA'}(\alpha, \alpha')} &= n \delta^{B'_n A} \delta^{A' B_2} \delta^{B'_2 B_3} \dots \delta^{B'_{n-1} B_n} \\ &\quad \times \delta(\beta'_n, \alpha) \delta(\alpha', \beta_2) \delta(\beta'_2, \beta_3) \dots \delta(\beta'_{n-1}, \beta_n) \\ &\quad \times \rho_{B_2 B'_2}(\beta_2, \beta'_2) \dots \rho_{B_n B'_n}(\beta_n, \beta'_n). \end{aligned} \quad (63)$$

The first of these implies the known result

$$[F, G, S_2] = -i \text{Tr}(\rho[\hat{F}, \hat{G}]) \quad (64)$$

leading to the von Neumann/Heisenberg equations for states/observables in linear QM

$$\frac{d}{d\tau} \text{Tr}(\rho \hat{F}) = -i \text{Tr}(\rho[\hat{F}, \hat{H}]). \quad (65)$$

The same equation is valid in the Polchinski-Jordan density matrix formulation of Weinberg's NLQM [18,13], but then $\hat{F} = \hat{F}[\rho]$, etc. Consider now a functional S (differentiable in f_k)

$$S[\rho] = S(f_1[\rho], \dots, f_n[\rho], \dots) \quad (66)$$

where $f_k[\rho] = \text{Tr}(\rho^k)$.

Theorem 3 *For any $m \in \mathbf{N}$, and any G , if S satisfies (66) then*

$$[f_m, G, S] = 0. \quad (67)$$

Proof:

$$[\text{Tr}(\rho^m), G, S] = \sum_n [\text{Tr}(\rho^m), G, f_n] \frac{\partial S}{\partial f_n} = -im \sum_n n \text{Tr}(\hat{G}[\rho^{m-1}, \rho^{n-1}]) \frac{\partial S}{\partial f_n} = 0. \quad (68)$$

□ This interesting result covers many nontrivial generalizations of S_2 . As a by-product it shows also that the same property holds for the Weinberg-Polchinski-Jordan NLQM because we have not assumed that G is linear in ρ (moreover, it includes other theories where observables do not satisfy any homogeneity condition). The particular case $m = 1$ implies that $\text{Tr} \rho$ is conserved by all evolutions, a fact important for a definition of averages. For pure states $\text{Tr}(\rho^m) = (\text{Tr} \rho)^m$ so that the integrals f_m are not necessarily independent, but for all m, n f_m and f_n are in involution with respect to $\{\cdot, \cdot\}_S$. Jordan proved in [13] by an explicit calculation that in his formulation of Weinberg's nonlinear QM $\text{Tr} \rho$ and $\text{Tr} \rho^2$ are conserved — our theorem considerably generalizes this result.

Theorem 4 *Let S satisfy (66) and ρ_t be a self-adjoint solution of (57). If ρ_0 is positive and has a finite number of nonvanishing eigenvalues $p_k(0)$, $0 < p_k(0) \leq 1$, then the eigenvalues of ρ_t are integrals of motion, and the evolution conserves positivity of ρ_t .*

Proof: Since the nonvanishing eigenvalues of ρ_0 satisfy $0 < p_k(0) \leq 1 < 2$, it follows that for any α $p_k(0)^\alpha$ can be written in a form of a convergent Taylor series. By virtue of the spectral theorem the same holds for ρ_0^α and $\text{Tr}(\rho_0^\alpha)$. Each element of the Taylor expansion of $\text{Tr}(\rho_0^\alpha)$ is proportional to $f_n[\rho_0]$, for some n . But $f_n[\rho_0] = f_n[\rho_t]$ hence

$$\text{Tr}(\rho_0^\alpha) = \text{Tr}(\rho_t^\alpha) = \sum_k p_k(0)^\alpha = \sum_k p_k(t)^\alpha \quad (69)$$

for all real α . Since all $p_k(0)$ are assumed to be known (the initial condition), we know also $\sum_k p_k(0)^\alpha = \sum_k p_k(t)^\alpha$ for any α . We can now apply the result used in the information theory [9] stating that the knowledge of $\sum_k p_k(t)^\alpha$ for all α uniquely determines $p_k(t)$. The continuity in t implies that $p_k(t) = p_k(0)$. □

The spectral decomposition of the density matrix

$$\rho_t = \sum_k p_k |k, t\rangle \langle k, t|, \quad (70)$$

where $t \mapsto |k, t\rangle$ defines a one-parameter continuous family of orthonormal vectors, leads to the unitary (although ρ -dependent) transformation $|k, t\rangle = U(\rho_t, \rho_0)|k, 0\rangle$. The density matrix evolves then as follows

$$\rho_t = U(\rho_t, \rho_0)\rho_0 U(\rho_t, \rho_0)^{-1}. \quad (71)$$

The question whether the same holds good for ρ_0 having an infinite number of nonvanishing eigenvalues will be left open here. In any case, it seems that the above theorem is sufficient at least “for all practical purposes”.

To make our proposal more concrete, we have to choose some explicit “physical” class of S — and here the information theoretic introduction may be helpful.

The suggestion of Wigner that a natural arena for nonlinear generalizations of the linear formalism of QM is the domain of *observations* leads to investigation of systems that *can gain information* hence are described by $\alpha \neq 2$ entropies. A homogeneity preserving generalization of S_2 for other α -entropies can be, for instance,

$$S_\alpha[\rho] = \left(1 - \frac{1}{\alpha}\right) \frac{(\text{Tr}(\rho^\alpha))^{1/(\alpha-1)}}{(\text{Tr} \rho)^{1/(\alpha-1)-1}}. \quad (72)$$

The choice of the denominator is important only from the point of view of the homogeneity of the evolution equation. The multiplier $1 - 1/\alpha$ guarantees that the evolution of pure states is the same, hence *linear*, for all α (this is reasonable as pure states have the same, vanishing α -entropies). The generalized Liouville-von Neumann equation following from (72) is

$$i \frac{d}{d\tau} \rho = \frac{(\text{Tr}(\rho^\alpha))^{1/(\alpha-1)-1}}{(\text{Tr} \rho)^{1/(\alpha-1)-1}} [\hat{H}, \rho^{\alpha-1}]. \quad (73)$$

For pure states and $\text{Tr} \rho = 1$, $\rho^n = \rho$ and the equation reduces to the ordinary, linear one; for mixed states the evolution is nonlinear unless the states are “so mixed” that ρ is proportional to the unit operator (which makes sense in finite dimensional cases, of course) and all α -entropies reduce to the Hartley formula.

The evolution of (now linear) observables is governed by

$$i \frac{d}{d\tau} F = \frac{(\text{Tr}(\rho^\alpha))^{1/(\alpha-1)-1}}{(\text{Tr} \rho)^{1/(\alpha-1)-1}} \text{Tr}(\rho^{\alpha-1} [\hat{F}, \hat{H}]) \quad (74)$$

which shows that for the generalized S the time derivative of an observable is not linear in the density matrix. For $\alpha = 2$ the equations reduce again to the ordinary linear equations.

It seems that the following choice of S_α is also interesting:

$$S_\alpha[\rho] = \frac{1}{2} \frac{(\text{Tr}(\rho^\alpha))^{1/(\alpha-1)}}{(\text{Tr} \rho)^{1/(\alpha-1)-1}}. \quad (75)$$

For pure states the expression reduces to the linear form $\frac{1}{2} \langle \psi | \psi \rangle^2 = \frac{1}{2} \text{Tr}(\rho^2)$. The density matrix would satisfy then the equation

$$i \frac{d}{d\tau} \rho = \frac{1}{2} \frac{\alpha}{\alpha - 1} \frac{(\text{Tr}(\rho^\alpha))^{1/(\alpha-1)-1}}{(\text{Tr} \rho)^{1/(\alpha-1)-1}} [\hat{H}, \rho^{\alpha-1}] \quad (76)$$

which for pure states and normalized ρ would become

$$2 \frac{\alpha - 1}{\alpha} i \frac{d}{d\tau} \rho = [\hat{H}, \rho] \quad (77)$$

and the “Boltzmann-Shannon classical limit” $\alpha \rightarrow 1$ of the Rényi entropy is indistinguishable from the $\hbar \rightarrow 0$ classical limit of QM.

D. Composition Problem for Subsystems with Different Entropies

Assuming that the formalism is applicable to a description of the composite “object+observer” system, where the nonlinearity is a feature of the *observer*, we have to know how to combine systems that are described by different entropies.

I think it is best to approach the question again in an information theoretic way. To begin with, let us consider a system whose entropy is I_α , and whose subsystems have entropies of the *same* kind: for all k a k -th system’s entropy satisfies $I_{\alpha_k} = I_\alpha$. Let the k -th subsystem be described by a reduced density matrix ρ_k . The entropy of the “large” system should be defined, as usual in information theory, as the average entropy of the subsystems. The overall entropy of the large systems should not depend on the way we decompose it into subsystems. Therefore the average cannot have the apparently natural form

$$I_{\alpha_1 \dots \alpha_n}[\rho] = \sum_k^n p_k I_\alpha[\rho_k], \quad (78)$$

where p_k are some weights, because the LHS is sensitive to correlations between the subsystems whereas the RHS is not, so that the entropy would be sensitive to the decompositions which are arbitrary. It seems we have to assume that in such a case the composition takes the trivial form

$$I_{\alpha_1 \dots \alpha_n}[\rho] = \sum_k p_k I_{\alpha_k}[\rho] = \sum_k p_k I_\alpha[\rho] = I_\alpha[\rho]. \quad (79)$$

Consider now a situation where the different subsystems have *different* entropies, say, I_{α_k} . The average entropy of the composite system is now defined in analogy to (79) as

$$I_{\alpha_1 \dots \alpha_n}[\rho] = \sum_{\alpha_k} p_k I_{\alpha_k}[\rho] \quad (80)$$

where the probabilities p_k are weights describing the “percentage” of each of the entropies in the overall entropy of the system. We do not know how to determine the weights — they can play a role of parameters characterizing the system.

The above definitions imply that the α^* -entropy of the large system is

$$I_{\alpha_1 \dots \alpha_n}^*[\rho] = \prod_{\alpha_k} I_{\alpha_k}^*[\rho]^{p_k}, \quad (81)$$

so it is natural to define

$$S_{\alpha_1 \dots \alpha_n}[\rho] = \prod_{\alpha_k} S_{\alpha_k}[\rho]^{p_k}. \quad (82)$$

Denoting the latter expression by S , we obtain

$$\frac{d}{d\tau} \rho_b = \sum_{\alpha_k} p_k [\rho_b, H, S_{\alpha_k}] \frac{S}{S_{\alpha_k}}. \quad (83)$$

Consider again a system which consists of subsystems equipped with the entropy of the same kind. Then $S_{\alpha_k} = S_{\alpha_l} = S$ for all k and l and the system evolves according to

$$\frac{d}{d\tau} \rho_b = [\rho_b, H, S] \quad (84)$$

as expected.

The next possibility is that the entropies that sum to the overall entropy are again sums of some other entropies. The description of the whole system should not depend on the order in which the partial entropies are summed up. So consider two entropies S^I and S^{II} , with appropriate weights λ^I and λ^{II} , and let the entropies S^I and S^{II} consist of some other entropies S_k^I and S_l^{II} appearing with weights $\{p_k^I\}_{k=1}^N$ and $\{p_l^{II}\}_{l=1}^M$, respectively. Then

$$\frac{d}{d\tau} \rho_b = \sum_k \lambda^I p_k^I \{\rho_b, H\}_{S_k^I} \frac{S}{S_k^I} + \sum_l \lambda^{II} p_l^{II} \{\rho_b, H\}_{S_l^{II}} \frac{S}{S_l^{II}} \quad (85)$$

which shows that the evolution can be indeed consistently composed of “sub-entropies”.

If all $S_{\alpha_k}[\rho]$ depend on ρ only *via* $f_m[\rho]$, like in our definitions (72) and (76), we know that on general grounds they are integrals of motion. Consider a subsystem described by ρ_k and which is noninteracting with the subsystem “where the nonlinearity resides”. In such a case the overall Hamiltonian function is $H[\rho] = \sum_k H_k[\rho_k]$ and

$$\frac{d}{d\tau}\rho_{kb} = \sum_l p_l[\rho_{kb}, H_k[\rho_k], S_{\alpha_l}] \frac{S}{S_{\alpha_l}}. \quad (86)$$

A system described by ρ_k and S_{α_k} can be totally isolated from the “rest of the Universe”, if for $k \neq l$, $[\rho_{kb}, H_k[\rho_k], S_{\alpha_l}] = 0$ and $p_k = S_{\alpha_k}/S$. However, even in such a case the global properties of the large system leave their mark on the local properties of all the subsystems as $p_k S/S_{\alpha_k}$ is at most an integral of motion hence depends on initial conditions. Consider a general H (including the interaction) and let

$$S[\rho] = S_{\alpha_1 \dots \alpha_n}[\rho] = \prod_{\alpha_k} S_{\alpha_k}[\rho_k]^{p_k}. \quad (87)$$

where the different subsystems have different entropies. Each of the sub-entropies satisfies

$$\frac{d}{d\tau} S_{\alpha_k}[\rho_k] = \sum_{\alpha_l} p_l[S_{\alpha_k}[\rho_k], H, S_{\alpha_l}(\rho_l)] \frac{S}{S_{\alpha_l}} = 0 \quad (88)$$

in virtue of the theorem 2 (we have used here the fact that $\{F, H\}_S = -\{F, S\}_H$, and the antisymmetry of the triple bracket. Therefore not only the overall entropy, but also the sub-entropies are integrals of motion *even if the Hamiltonian function H contains interaction terms*).

Consider now the reduced density matrix ρ_k of one of the subsystems. Using the same theorem we find that

$$\frac{d}{d\tau}\rho_{kb} = p_k[\rho_{kb}, H, S_{\alpha_k}] \frac{S}{S_{\alpha_k}}. \quad (89)$$

where S/S_{α_k} is an integral of motion but its value depends on initial conditions. If the change of the initial conditions does not affect the reduced density matrices in (87), the integral of motion is also unchanged. Therefore in order to change this quantity we have to change correlations between the subsystems. In particular, a time dependence of a linear system which is noninteracting with the nonlinear one is insensitive to changes of initial conditions within the linear system if the particular form (87) holds. For global entropies different from (87) some kind of sensitivity appears but the influences between the subsystems cannot propagate faster than light unless we introduce the projection postulate.

Such a trace of nonlinearity observed in some linear system might be used to detect the nonlinearity. Following Santilli [19] we can expect that an evolution of an internal part of a hadron may be nonlinear (like in hadronic mechanics). In such a case correlations between a hadron (say, a proton) and some linear system (say, an electron) could be observed in a form of a ρ -dependent rescaling of time in the electron’s evolution.

V. COMMENTS

S. Weinberg wrote in [20] that the “theoretical failure to find a plausible alternative to quantum mechanics, even more than the precise experimental verification of linearity, suggests (...) that quantum mechanics is the way it is because any small change in quantum mechanics would lead to logical absurdities. If this is true, quantum mechanics may be a permanent part of physics. Indeed, quantum mechanics may survive not merely as an approximation to a deeper truth, (...) but as a precisely valid feature of the final theory.” This kind of conviction followed from the internal theoretical difficulties of the generalizations based on nonlinear Schrödinger equations and general Hamiltonian framework. These difficulties have been discussed in detail in [1].

The proposal based on the generalized Nambu dynamics is free from those difficulties. However, it has new features with respect to ordinary QM. One, the stronger complementarity principle, has already been announced. In the generalized framework a time derivative of an observable will not, in general, be linear in the density matrix. Since we have defined observables as functions necessarily linear in ρ , the time derivatives of observables are not themselves observables.

I propose the following interpretation of this fact. To focus our attention let us consider linear QM and the nonrelativistic position operator. An average velocity of an ensemble of particles can be calculated either by first

calculating an average position and then taking its time derivative, or by first measuring the velocity of each single particle and then taking the average. We can say that the first procedure is a calculation of the time derivative of an average, whereas the latter is taking the average of the time derivative. The situation can be described symbolically by the equation

$$\frac{d}{dt}\langle\vec{q}\rangle = \langle\frac{d}{dt}\vec{q}\rangle. \quad (90)$$

An important property of QM is the impossibility of realizing the two procedures simultaneously, as $\vec{v} = \vec{p}/m$ and \vec{q} are *complementary*. It follows that in a concrete experiment we have to decide which way of measuring to choose. In this meaning if we can measure \vec{q} , we cannot measure $\frac{d}{dt}\vec{q}$, and *vice versa*. To express it differently, if \vec{q} is observable (not *an* observable!) then $\frac{d}{dt}\vec{q}$ is not.

In triple bracket NLQM the observables will be defined as quantities that are in one-to-one relationship to some experimentally measured random variables (hence the linearity in ρ). Two observables will be said to be complementary if there does not exist a physical situation where the two respective random variables can be measured *simultaneously*, that is, in a single run of an experiment. There can exist linear operators representing the position and the velocity of single members of an ensemble, but if the ensemble evolves in a nonlinear way, the averages of those observables do not have to satisfy the inherently linear condition (90), if the experimental procedures necessary for their measurements cannot be simultaneously realized.

Another fundamental problem, arising in the Nambu-like description, is the action principle leading to the triple bracket equation. The Hamiltonian NLQM proposed by Kibble or Weinberg can be derived from the ordinary Lagrangian formalism. The triple bracket form of dynamics must follow from a new kind of variational principle. The variational principle proposed recently by Takhtajan [21] suggests an interesting direction for further investigations.

VI. APPENDIX: HAMILTONIAN FORMULATION OF THE DIRAC EQUATION

Consider the Dirac equation

$$(i\gamma^a\nabla_a - m)\psi = 0 \quad (91)$$

where $\nabla_a = \partial_a + ie\Phi_a$ and Φ_a is an electromagnetic potential world-vector. We are going to rewrite the equation in a form of the “proper-time” covariant Hamilton equations of motion. The “proper time” will be defined in terms of spacelike hyperplanes constructed as follows. Let $\sigma_\tau(x(\tau)) = 0$ be an equation defining a family of spacelike hyperplanes. The field of timelike, future-pointing, normalized vectors $n_\tau^a(x) \propto \partial^a\sigma_\tau(x)$, satisfying the continuity equation $\partial_a n_\tau^a(x) = 0$, defines the field of “proper time” directions. Integral curves $\tau \mapsto x^a(\tau)$ of $n_\tau^a(x)$, where τ is the parameter of the family $\{\sigma_\tau\}$, play the role of the world-lines. We shall need the continuity equation to guarantee the reality of the Hamiltonian function. Notice that this condition eliminates some physically meaningful hyperplanes, like the proper-time hyperboloid $\sigma_\tau(x) = x^a x_a - \tau^2 = 0$, but admits simultaneity hyperplanes $\sigma_\tau(x) = n^a x_a - \tau = 0$. The “proper time” following from the construction should not, for this reason, be identified with the ordinary proper time of the electron. The “proper time” derivative at x is defined as

$$\frac{d}{d\tau} = n_\tau^a(x)\partial_a. \quad (92)$$

Multiplying (91) from left by the Dirac matrices we obtain [22]

$$(i\nabla^a + \sigma^{ab}\nabla_b - m\gamma^a)\psi = 0. \quad (93)$$

Writing the four-potential explicitly in

$$i\partial^a\psi = (-\sigma^{ab}\partial_b + e\Phi^a + ie\sigma^{ab}\Phi_b + m\gamma^a)\psi \quad (94)$$

and contracting with $n_\tau^a(x)$ we get

$$\begin{aligned} i\frac{d}{d\tau}\psi(x) &= \left(-\sigma_{ab}n_\tau^a(x)\partial^b + en_\tau^a(x)\Phi_a(x) + ie\sigma_{ab}n_\tau^a(x)\Phi^b(x) \right. \\ &\quad \left. + m n_\tau^a(x)\gamma_a \right)\psi(x) \end{aligned} \quad (95)$$

$$= \hat{H}\psi(x). \quad (96)$$

In the spinor language

$$i\nabla^{AA'}\phi_A = \mu\chi^{A'} = ig_a^{AA'}\nabla^a\phi_A \quad (97)$$

$$i\nabla_{AA'}\chi^{A'} = -\mu\phi_A = ig_{aAA'}\nabla^a\chi^{A'} \quad (98)$$

where $\mu = m/\sqrt{2}$ and $g_a^{AA'}$ are the Infeld-van der Waerden symbols [12]. Using the identities

$$g^a_{XA'}g^{bYA'} + g^b_{XA'}g^{aYA'} = g^{ab}\varepsilon_X^Y \quad (99)$$

$$g^a_{XA'}g^{bYA'} - g^b_{XA'}g^{aYA'} = 4\sigma^{ab}_X{}^Y \quad (100)$$

$$g^a_{AX'}g^{bAY'} - g^b_{AX'}g^{aAY'} = 4\bar{\sigma}^{ab}_{X'}{}^{Y'} \quad (101)$$

where $\sigma^{ab}_X{}^Y$ and $\bar{\sigma}^{ab}_{X'}{}^{Y'}$ are generators of $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations of $SL(2, \mathbf{C})$ we obtain

$$i\nabla_a\phi_X = -4i\sigma_{abX}{}^Y\nabla^b\phi_Y + 2\mu g_{aXX'}\chi^{X'} \quad (102)$$

$$i\nabla_a\chi^{X'} = 4i\bar{\sigma}_{abY'}{}^{X'}\nabla^b\chi^{Y'} - 2\mu g_a^{XX'}\phi_X \quad (103)$$

and the equations obtained by their complex conjugation. These equations are especially simple if we express generators and Infeld-van der Waerden symbols in purely spinorial terms. Remembering that $n_\tau^a n_{a\tau} = 1$ implies $n_{AA'\tau} n_\tau^{BA'} = \frac{1}{2}\varepsilon_A^B$ we get after some calculations

$$\begin{aligned} i\frac{d}{d\tau}\phi_X(x) &= i n_\tau^{YY'}(x)\nabla_{XY'}\phi_Y(x) + \mu n_{\tau XX'}(x)\chi^{X'}(x) \\ &\quad + e n_\tau^a(x)\Phi_a(x)\phi_X(x) \end{aligned} \quad (104)$$

$$\begin{aligned} i\frac{d}{d\tau}\chi^{X'}(x) &= i n_{\tau YY'}(x)\nabla^{YX'}\chi^{Y'}(x) - \mu n_\tau^{XX'}(x)\phi_X(x) \\ &\quad + e n_\tau^a(x)\Phi_a(x)\chi^{X'}(x) \end{aligned} \quad (105)$$

Let $d\sigma_\tau(x)$ be some invariant measure on the hyperplane σ_τ . The equations can be derived from the Hamiltonian function

$$\begin{aligned} H[\psi, \psi^*] &= \langle \psi | \hat{H} | \psi \rangle \\ &= \int_{\sigma_\tau} \left\{ i\phi_{X'}^*(x)n_\tau^{XX'}(x)n_\tau^{YY'}(x)\nabla_{XY'}\phi_Y(x) \right. \\ &\quad \left. - i\bar{\chi}^{*X}(x)n_{\tau XX'}(x)n_{\tau YY'}(x)\nabla^{YX'}\chi^{Y'}(x) \right. \\ &\quad \left. + \frac{1}{2}\mu \left(\phi_{X'}^*(x)\chi^{X'}(x) + \chi^{*X}(x)\phi_X(x) \right) \right. \\ &\quad \left. + e n_\tau^a(x)\Phi_a(x)n_\tau^{XX'}(x) \left(\phi_X(x)\phi_{X'}^*(x) + \chi_X^*(x)\chi_{X'}(x) \right) \right\} d\sigma_\tau(x) \end{aligned} \quad (106)$$

provided

$$\partial^{YX'}n_{\tau XX'}(x) = \partial^{XX'}n_{\tau XY'}(x) = 0 \quad (107)$$

and the wave functions vanish at boundaries of the hyperplane σ_τ . Reality of H is guaranteed by the same conditions. The Hamiltonian function is not positive definite, which is correct since we are working here in first quantized formalism. Contraction of (107) over the remaining indices implies the continuity equation discussed above.

The explicit form of the Hamilton equations is

$$i n_\tau^{XX'}(x)\frac{d}{d\tau}\phi_X(x) = \frac{\delta H}{\delta \phi_{X'}^*(x)}, \quad (108)$$

$$i n_{\tau XX'}(x)\frac{d}{d\tau}\chi^{X'}(x) = \frac{\delta H}{\delta \chi^{*X}(x)}, \quad (109)$$

and c.c, or, in the Poissonian way,

$$i \frac{d}{d\tau} \phi_X(x) = 2 n_{\tau X X'}(x) \frac{\delta H}{\delta \phi_{X'}^*(x)}, \quad (110)$$

$$i \frac{d}{d\tau} \chi^{X'}(x) = 2 n_{\tau}^{X X'}(x) \frac{\delta H}{\delta \chi^{*X}(x)}, \quad (111)$$

We can see that $i n_{\tau X X'}(x) = \omega_{\tau X X'}(x)$ are the components of the symplectic (since derivable from a Kähler potential $\|\psi\|^2$) form on σ_τ at point $x \in \sigma_\tau$, and the Poissonian form $I_{\tau X X'}(x) = -2i n_{\tau X X'}(x)$.

Let $\gamma_a^{\alpha\beta}$, $a = 0, 1, 2, 3$, be the Dirac matrices. The Hamilton equations equivalent to the Dirac equation written in the bispinor form are

$$i n_\tau^a(x) \gamma_a^{\alpha\beta} \frac{d}{d\tau} \psi_\beta(x) = \frac{\delta H}{\delta \psi_\alpha^*(x)}, \quad (112)$$

and c.c., where $*$ denotes the complex conjugation. The formulas are simplest if we take simultaneity hyperplanes foliation of the Minkowski space. Then $n_\tau^a(x) \gamma_a^{\alpha\beta} = \gamma_0^{\alpha\beta}$. Denoting its inverse by $\gamma_{0\alpha\beta}$ we find that

$$\frac{d}{d\tau} \psi_\alpha(x) = -i \gamma_{0\alpha\beta} \frac{\delta H}{\delta \psi_\beta^*(x)} \quad (113)$$

$$\frac{d}{d\tau} \psi_\alpha^*(x) = i \gamma_{0\alpha\beta} \frac{\delta H}{\delta \psi_\beta(x)} \quad (114)$$

So here the Poissonian form is $I_{\alpha\beta} = -i \gamma_{0\alpha\beta}$ and the Dirac matrix $\gamma_{0\alpha\beta}$ corresponds to $\delta_{AB'}$ discussed in III. The transition to the triple bracket formalism is now straightforward.

* Address after 15 August 1994: Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307

Electronic address: mczachor@sunrise.pg.gda.pl

- [1] M. Czachor, *Elements of nonlinear quantum mechanics (I): Nonlinear Schrödinger equation and two-level atoms*, submitted to Phys. Rev. A.
- [2] I. Białynicki-Birula and P. J. Morrison, Phys. Lett. A **158**, 453 (1991).
- [3] Y. Nambu, Phys. Rev. D **7**, 2405 (1973).
- [4] A. O. Barut and R. Rączka, *Theory of Group Representations and Applications*, (Polish Scientific Publishers, Warszawa, 1980).
- [5] E. P. Wigner, *Symmetries and Reflections*, (Indiana University Press, Bloomington, 1967).
- [6] R. V. Hartley, Bell Syst. Tech. Jour. **7**, 535 (1928).
- [7] C. E. Shannon, Bell Syst. Tech. Jour. **27**, 379 (1948).
- [8] N. Wiener, *Cybernetics*, Wiley, New York (1948).
- [9] A. Rényi, in *Selected Papers of Alfréd Rényi*, Akadémiai Kiadó, Budapest (1976).
- [10] Z. Daróczy, Inf. Control **16**, 36 (1970).
- [11] U. Fano, Rev. Mod. Phys. **29**, 74 (1957).
- [12] R. Penrose and W. Rindler, *Spinors and Space-Time*, Vol. 1, Cambridge University Press, Cambridge (1984).
- [13] T. F. Jordan, "Reconstructing a Nonlinear Dynamical Framework for Testing Quantum Mechanics", preprint (1991).
- [14] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, Berlin (1989).
- [15] A. P. Fordy, in *Nonlinear Fields: Classical, Random, Semiclassical*, edited by P. Garbaczewski and Z. Popowicz, World Scientific, Singapore (1991).
- [16] T. W. Kibble, Comm. Math. Phys. **64**, 73 (1978). *ibid.* **65**, 189 (1979).
- [17] S. Weinberg, *Ann. Phys.* (NY) **194**, 336 (1989).
- [18] J. Polchinski, Phys. Rev. Lett. **66**, 397 (1991).
- [19] R. M. Santilli, *Foundations of Theoretical Mechanics*, Springer-Verlag, New York (1982).
- [20] S. Weinberg, *Dreams of a Final Theory*, Hutchinson (1993).
- [21] L. Takhtajan, Comm. Math. Phys. **160**, 295 (1994).
- [22] S. Hacyan, Gen. Rel. Grav. **26**, 85 (1994).